

Лекция 9

МНОГОФУНКЦИОНАЛЬНЫЕ СТАТИСТИЧЕСКИЕ КРИТЕРИИ

Понятие многофункциональных критериев

Многофункциональные статистические критерии - это критерии, которые могут использоваться по отношению к самым разнообразным данным, выборкам и задачам.

Это означает, что данные могут быть представлены в любой шкале, начиная от номинативной

**Критерий φ^* — угловое
преобразование Фишера**

Назначение критерия

φ^*

Критерий Фишера предназначен для сопоставления двух выборок по частоте встречаемости интересующего исследователя эффекта.

Описание критерия

Критерий оценивает достоверность различий между процентными долями двух выборок, в которых зарегистрирован интересующий нас эффект.

Суть углового преобразования Фишера состоит в переводе процентных долей в величины центрального угла, который измеряется в радианах.

Большей процентной доле будет соответствовать больший угол ϕ , а меньшей доле - меньший угол, но соотношения здесь не линейные:

$$\phi = 2 \cdot \arcsin(\sqrt{P})$$

где P - процентная доля, выраженная в долях единицы

Гипотезы

H_0 : Доля лиц, у которых проявляется исследуемый эффект, в выборке 1 не больше, чем в выборке 2.

H_1 : Доля лиц, у которых проявляется исследуемый эффект, в выборке 1 больше, чем в выборке 2.

Ограничения критерия φ^*

1. Ни одна из сопоставляемых долей не должна быть равной нулю. Формально нет препятствий для применения метода φ в случаях, когда доля наблюдений в одной из выборок равна 0. Однако в этих случаях результат может оказаться неоправданно завышенным
2. Верхний предел в критерии φ отсутствует - выборки могут быть сколь угодно большими. Нижний предел - 2 наблюдения в одной из выборок. Однако должны соблюдаться следующие соотношения в численности двух выборок:

а) если в одной выборке всего 2 наблюдения, то во второй должно быть не менее 30:

$$n_1=2 \rightarrow n_2 \geq 30;$$

б) если в одной из выборок всего 3 наблюдения, то во второй должно быть не менее 7:

$$n_1=3 \rightarrow n_2 \geq 7;$$

в) если в одной из выборок всего 4 наблюдения, то во второй должно быть не менее 5:

$$n_1=4 \rightarrow n_2 \geq 5;$$

г) при $n_1, n_2 \geq 5$ возможны любые сопоставления.

В принципе возможно и сопоставление выборок,

не отвечающих этому условию, например, с

соотношением $n_1=2, n_2=15$, но в этих случаях не

удастся выявить достоверных различий.

АЛГОРИТМ

Расчет критерия ϕ^*

1. Определить те значения признака, которые будут критерием для разделения испытуемых на тех, у кого "есть эффект" и тех, у кого "нет эффекта".
2. Начертить четырехклеточную таблицу из двух столбцов и двух строк. Первый столбец - "есть эффект"; второй столбец - "нет эффекта"; первая строка сверху - 1 группа (выборка); вторая строка - 2 группа (выборка).
3. Подсчитать количество испытуемых в первой группе, у которых "есть эффект", и занести это число в левую верхнюю ячейку таблицы.

4. Подсчитать количество испытуемых в первой выборке, у которых "нет эффекта", и занести это число в правую верхнюю ячейку таблицы.

Подсчитать сумму по двум верхним ячейкам. Она должна совпадать с количеством испытуемых в первой группе.

5. Подсчитать количество испытуемых во второй группе, у которых "есть эффект", и занести это число в левую нижнюю ячейку таблицы.

6. Подсчитать количество испытуемых во второй выборке, у которых "нет эффекта", и занести это число в правую нижнюю ячейку таблицы.

Подсчитать сумму по двум нижним ячейкам. Она должна совпадать с количеством испытуемых во второй группе (выборке).

7. Определить процентные доли испытуемых, у которых "есть эффект", путем отнесения их количества к общему количеству испытуемых в данной группе (выборке).

Записать полученные процентные доли соответственно в левой верхней и левой нижней ячейках таблицы в скобках, чтобы не перепутать их с абсолютными значениями.

8. Проверить, не равняется ли одна из сопоставляемых процентных долей нулю. Если это так, попробовать изменить это, сдвинув точку разделения групп в ту или иную сторону. Если это невозможно или нежелательно, отказаться от критерия ϕ^* и использовать критерий χ^2 .

9. Определить по Табл. XII Приложения 1 величины углов φ для каждой из сопоставляемых процентных долей.

10. Подсчитать эмпирическое значение φ^* по формуле:

$$\varphi^* = (\varphi_1 - \varphi_2) \cdot \sqrt{\frac{n_1 \cdot n_2}{n_1 + n_2}}$$

где: φ_1 - угол, соответствующий большей процентной доле;

φ_2 - угол, соответствующий меньшей процентной доле;

n_1 - количество наблюдений в выборке 1;

n_2 - количество наблюдений в выборке 2.

11. Сопоставить полученное значение φ^* с критическими значениями: $\varphi^* \leq 1,64$ ($P < 0,05$)
И $\varphi^* \leq 2,31$ ($p < 0,01$).

Если $\varphi^*_{\text{эмп}} \leq \varphi^*_{\text{кр}}$. H_0 отвергается.

При необходимости определить точный уровень значимости полученного $\varphi^*_{\text{эмп}}$ по Табл. XIII Приложения 1.

Пример 1

сопоставление выборок по качественно определяемому признаку

В данном варианте использования критерия мы сравниваем процент испытуемых в одной выборке, характеризующихся каким-либо качеством, с процентом испытуемых в другой выборке, характеризующихся тем же качеством.

Допустим, нас интересует, различаются ли две группы студентов по успешности решения новой экспериментальной задачи. В первой группе из 20 человек с нею справились 12 человек, а во второй выборке из 25 человек - 10. В первом случае процентная доля решивших задачу составит $12/20 \cdot 100\% = 60\%$, а во второй $10/25 \cdot 100\% = 40\%$. Достоверно ли различаются эти процентные доли при данных n_1 и n_2 ?

Казалось бы, и "на глаз" можно определить, что 60% значительно выше 40%. Однако на самом деле эти различия при данных n_1 , n_2 недостоверны.

Проверим это. Поскольку нас интересует факт решения задачи, будем считать "эффектом" успех в решении экспериментальной задачи, а отсутствием эффекта - неудачу в ее решении.

Сформулируем гипотезы.

H_0 : Доля лиц, справившихся с задачей, в первой группе не больше, чем во второй группе.

H_1 : Доля лиц, справившихся с задачей, в первой группе больше, чем во второй группе.

Четырехклеточная таблица для расчета критерия при сопоставлении двух групп испытуемых по процентной доле решивших задачу.

группы	Есть эффект решение задачи			Нет эффекта			суммы
	Ко-во	%-доля		Кол-во	%-доля		
1 груп	12	60%	А	8	40%	Б	20
2 груп	10	40%	В	15	60%	Г	25
Сумма	22			23			45

В четырехклеточной таблице, как правило, сверху размечаются столбцы "Есть эффект" и "Нет эффекта", а слева - строки "1 группа" и "2 группа". Участвуют в сопоставлениях, собственно, только поля (ячейки) А и В, то есть процентные доли по столбцу "Есть эффект".

По Табл. XII Приложения 1 определяем величины φ , соответствующие процентным долям в каждой из групп.

$$\varphi_{1(60\%)} = 1,772$$
$$\varphi_{2(40\%)} = 1,369$$

Теперь подсчитаем эмпирическое значение φ^*
по формуле:
В данном случае:

$$\varphi^*_{\text{эм}} = (1,772 - 1,369) \cdot \sqrt{\frac{20 \cdot 25}{20 + 25}} = 0,403 \cdot \sqrt{11,11} = 1,34$$

По Табл. XIII Приложения 1 определяем,
какому уровню значимости соответствует

$$\varphi^*_{\text{эмп}}=1,34:$$

$$p=0,09$$

Можно установить и критические значения φ^* , соответствующие принятым в психологии уровням статистической значимости:

$$\varphi_{\text{кр}}^* = \begin{cases} 1,64 (p \leq 0,05) \\ 2,31 (p \leq 0,01) \end{cases}$$

$$\varphi_{\text{вып}} = 1,34$$

$$\varphi_{\text{вып}} < \varphi_{\text{кр}}$$

Полученное эмпирическое значение φ^*
находится в зоне незначимости.

Ответ: H_0 принимается. Доля лиц,
справившихся с задачей, в первой группе не
больше, чем во второй группе.